DOCUMENT RESUME

ED 319 778 TM 014 996

AUTHOR Lambert, Nadine M.
TITLE The Crisis in Measurement Literacy in Psychology and Education.
PUB DATE Aug 89
NOTE 16p.; Paper presented at the Annual Meeting of the American Psychological Association (97th, New Orleans, LA, August 11-15, 1989).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Accreditation (Institutions); Curriculum Development; *Educational Researchers; *Measurement Techniques; National Programs; *Psychologists; *Psychometrics; Qualifications; Specialists; State Licensing Boards; *Statistical Analysis; Testing Programs; *Test Interpretation

ABSTRACT

Issues associated with the lack of statistical skills among educational and psychological researchers are discussed. Issues are associated with accreditation standards, faculty positions for psychometric specialists, measurement curricula, and licensing organizations. The Test User Qualification Working Group has produced a model test user qualification system to serve as a tool for identifying competent test users regardless of educational background or title. The group identified 86 generic subelements or knowledge, skills, abilities, and other characteristics relevant to preventing misuse of the more than 50 sample commercial tests used in a test user study. Eight empirically derived clusters of tests were also identified. Other suggested aids for improving measurement literacy address instruction; methods for assessing program compliance with accreditation rules; participation in national conferences on graduate and continuing education; integration of measurement instruction into proposals for reform in teacher education; review of the content of national examinations for practitioners; establishment of state guidelines for school testing specialists; and creation of fellowships, academic chairs, and prizes for promising measurement students. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
NADINE M. LAMBERT
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED319778

The Crisis in Measurement Literacy in Psychology and Education [1]

Nadine M. Lambert

University of California, Berkeley

The topic of this address was prompted by many experiences and published commentary over the past decade that raise serious questions about our success as measurement specialists in translating our ever increasing sophistication in psychometric theory to the psychological and educational publics who use our various products, as well as the extent to which we have been successful in sustaining and nurturing instruction in measurement theory and in our many institutions of higher learning.

What is "measurement literacy"

To have achieved minimal literacy in measurement, as a psychologist, an educator, a personnel manager, a counselor, or other human service practitioner, one would have acquired at least four types of knowledge:

1) Knowledge of the basic assumptions that underlie the rendering or quantification of observations, assigning objects or events to classes, ordering of units of observations from greatest to smallest, or transforming of the number of right and wrong test answers on a formal or informal test.

2) Familiarity with the general rules by which observations, rank orders, item scores, and individual difference data are translated into

[1] Invited address presented at the Annual Convention of the American Psychological Association, New Orleans, LA, August 13, 1989.

BEST COPY AVAILABLE

TM014996

measurement units such as frequency counts, probability estimates, measures of central tendency, and measures of variability.

3) Measurement literacy implies familiarity with concepts of validity and reliability and the ability to utilize these concepts in selecting, using, or interpreting numbers that are derived from the several approaches to educational and psychological measurement.

4) And fourthly knowledge of sources of error and the ability to apply appropriate standard errors of measurement in making psychological or educational diagnoses, classifications, inferences, or predictions.

Is there a crisis in measurement literacy?

The crisis in measurement literacy is evident in several professional settings. Even though accreditation standards require courses in measurement, students often are not exposed to instruction in psychometric theory as part of their programs. Courses in individual testing, program evaluation, or statistics are assumed as satisfying the measurement standard. Psychologists on accreditation panels observe frequently, and express concern about the fact, that psychology and psychologists are losing their measurement foundation.

Members of departmental faculty who have argued for sustaining or increasing the number of faculty positions for psychometric specialists encounter considerable departmental resistance to their justifications of the importance of adding or retaining a position as a professor in measurement when few courses are offered or taken by undergraduate and graduate students in both psychology and education units. If departments don't have a measurement curriculum, they don't need psychometric faculty.

Officials on psychology licensure boards who review academic credentials of applicants for psychology licenses note that even though the licensure standards require at least some instruction in measurement theory, that such instruction is subsumed by courses in assessment or testing rather than courses in psychometric theory.

The perception of many people, including otherwise credible professionals in education and psychology, that the elimination of

standardized tests will eliminate bias in personnel selection and higher education selection, is a telling comment on the lack of clear thinking that is a necessary part of literacy

Those of us who are involved with educational uses of tests in elementary, secondary, or higher education typically assume incorrectly that teachers, school administrators, and other education professionals have sufficient knowledge to select and use tests that satisfy, at least minimally, the Standards for Educational and Psychological Tests.

The demand for measurement literacy

Test User Qualifications

We are indebted to the Test User Qualification Working Group for producing a model test user qualification system (Eyde, Moreland, Robertson, Primoff, and Most, 1988) to serve as a tool for identifying competent test users regardless of educational background or title. The current state of measurement literacy is evident from their analysis of the content domains of test user qualifications, and the types of test misuse for various types of tests.

They identified 86 generic subelements, or knowledges, skills, abilities, and other characteristics which were relevant to preventing misuse of the more than 50 sample commercial tests used in the test user study. Next they proceeded to identify both the minimum essentials or subelements of good test use and the more comprehensive requirements for engaging in good testing practices Minimum essentials included "keeping scoring keys and test materials under close scrutiny", and"not making photocopies of copyrighted materials," Examples of more sophisticated knowledge were the necessity to compare test scores with other data in a psychological history, being able to use the standard error of measurement, appropriate consideration of measurement error in using cutoff scores, and being able to select and use norm reference scores appropriately.

In order to guide test purchase by reference to user qualifications, 8 empirically derived clusters of tests were identified. These were classified as l) group educational tests, 2) ability and preference tests, 3) learning disability

and neuropsychological tests, 4) individual intelligence tests, 5) readiness tests, 6) objective personality tests, 7) self-administered and self-scored tests, and 8) projective tests. Their research, therefore, provides an excellent framework for setting standards for measurement literacy in the various professional practice domains where particular types of tests are used. Each of these clusters of tests can be matched with groups of psychological and educational professional specializations. For example, group educational tests are the province of the teacher and education professional. Objective and projective personality tests are used primarily by clinical and counselling psychologists, although I have read reports of the use of graphology in personnel selection.

The specification of test user competencies resulting from the efforts of the Test User Work Group would be especially useful to the APA Committee on Accreditation, as they evaluate the measurement competencies provided in a program's professional preparation curriculum. Moreover, state licensure boards and those who prepare examinations for licensure could use these findings to evaluate the evidence of measurement competence submitted by an individual who wishes to be licensed to practice independently and who will be a future test user. In turn, department chairs and faculty can use the test user competency model as a basis for developing and evaluating current and future measurement instruction.

Before the theoretically-oriented psychometrician argues that promoting test user competency is a problem to be solved by someone else, it is important to keep in mind that by providing guidelines for test user competence, and relating these guidelines to program requirements for accreditation or approval, one creates a need and increases the demand for measurement instruction in our colleges and universities. As the needs become evident, and demands for help increase, campus administrators and personnel committees will respond by securing faculty resources for measurement instruction, research, and theory development.

Multiple perspectives on teachers and testing

Lambert (1980) reported the results of a survey of representatives of teacher unions, deans of schools of education, and legislators regarding opinions of tests and testing. The most frequent answers from all three

groups of respondents about teacher attitudes toward tests and testing were "negative, afraid of results, suspicious, a threat to job security." The majority of deans and legislators felt that teachers' attitudes toward tests should be changed and they suggested that: 1) teachers learn more about existing tests and how to interpret them; 2) teachers accept tests as useful measures, and; 3) teachers not rely on tests as the sole source of information about performance differences.

Nearly all of the respondents to the survey stated that it was important for teachers to produce superior classroom tests. And responses to questions about criterion-referenced tests suggested that although these might provide alternatives to nationally standardized tests, teachers were judged not to understand the concepts underlying criterion-referenced measurement.

The results of the survey showed that demand for measurement literacy of teachers was generally recognized; however, of the 102 deans of schools of education who responded to the survey, only one quarter of them said that instruction in testing was offered in a 3-4 semester hour course. Another one third provided measurement training as a segment of another course, and one fourth of them stated that they had no intention of offering any instruction in measurement.

Although this is only one set of data on teacher attitudes about tests, and the need for teacher measurement competency, the results reflect widespread support for measurement literacy among teachers on the part of teacher organizational leaders, heads of teacher training institutions, and our legislators.

Measurement Literacy in the Practice of Psychology

The Committees appointed to draft the several documents providing policy and guidance for the use of tests in academic, applied, and professional psychology have worked diligently to produce documents that can be understood by a broad range of test users. But when APA Council Representatives and committee and board members, become involved in governance review of such documents as the Standards for Educational and Psychological Tests, (APA, 1985) the Code of Fair Testing Practices in Education, (Fremer, Diamond, and Camara, 1989) the Guidelines for Computer-Based Tests and Interpretation (APA, 1986), or the proposals for

test user guidelines, they soon recognized that the difficulty level of the concepts defined in the Standards far exceeded the understanding of many psychologists. To resolve this dilemma, the APA Council of Representatives appointed a special review committee composed of a variety of academic and professional psychologists who were themselves not measurement experts to review the final prepublication draft of the Standards to review and to propose language for the Standards that the typical test user could understand.

The experience of those reviewing the Standards provided eloquent testimony to the failure of our doctoral programs for insuring minimal competence in measurement among psychologists generally. Criteria for recognition of doctoral programs and doctoral preparation for professional practice for psychology have specified the need for basic instruction in measurement theory for many years. Both the Criteria for Accreditation (APA 1979) as well as standards for licensure (American Association of State and Psychology Boards and the Professional Examination Service, 1984) require knowledge of basic measurement principles as distinct from preparation in research methodology and research design. The examination for licensure candidates also includes extensive sections on measurement. One strategy for addressing measurement literacy among psychologists and educators is to analyze item performance data on measurement items on licensure and teacher competency examinations. As the item data from these examinations will now be made available, it should be possible to obtain empirical evidence on the proportion of licensure and credential candidates who perform at an acceptable level on questions pertaining to basic knowlege of measurement theory and technique. And if psychometricians are critical of the coverage of measurement content in the licensure examination, the various state licensure and credentialling boards, as well as the American Association of State Psychology Boards, would be responsive to suggestions as to how the content of the examination can be improved. And as requirements for measurement knowledge in preprofessional programs become more explicit and regulated, the demand for measurement instruction in psychology departments and professional psychology programs will follow suit.

The supply side of the measurement literacy question

Promoting a complement of measurement faculty and measurement courses in university programs. It seems pretty obvious that the number of faculty positions with measurement as a primary field of study, as well as the number of students interested in measurement as a career has been generally declining. As Davison, Damrin, and Drasgow (1986) concluded from their analysis of the number of doctoral programs in measurement and the number of graduate students, while the number of programs (66) remained relatively constant from 1973 to 1983, the proportion of programs has not matched the growth in psychology programs over the same period and dropped from .18 to .11. The recent study by Aiken, West, Sechrest, and Reno (1989) reported that one third of psychology departments offered no training in measurement. Although nearly all departments provided courses in statistics, only 17% of psychology departments offered a specialization in quantitative areas. Over all of the departments surveyed, only 108 students were enrolled in programs emphasizing quantitative methods, and of these only a portion were in the measurement area.

The Aiken et al. study forecasts the increasingly common challenge that will have to be faced when measurement faculty retire and department heads have to secure replacements. A case in point is the relatively recent retirement of a measurement instructor in my own institution. As head of the Division of Educational Psychology, I prepared several memoranda presenting arguments for retaining the position. The administration countered that there were few applicants for the measurement program, that measurement could be taught by statisticians in other departments, and that a professor holding the position would not be perceived as providing courses for which there would be a high student demand generally.

Faced with this challenge, I responded that the state of measurement literacy among educators was at an appallingly low level, and that schools of education aiming to forge new directions in the preparation of graduate students for service to schools and agencies required experts in measurement theory as much as they required experts in computer technology. Measurement theory, I argued, was a more fundamental area of instruction in graduate programs than was instruction in computer applications to

teaching. To counter the argument that other faculty could carry the measurement courses, I pointed out that psychometrics was a field of its own, and only an appropriately prepared psychometrician would be expected to have a research program that would support instruction in classical and modern test theory and their applications to common psychological and educational problems.

By pointing out that our success in recruiting students to our applicant pool was dependent on visible measurement faculty, as well as a high level measurement curriculum, I hoped to neutralize the argument with which I was faced, that student demand should be the criterion for faculty resource allocation. So it was necessary to argue for the wide applicability of measurement instruction to program evaluation, teacher preparation, preparation of psychologists to work in schools, as well as for many students in education whose own research required attention to measurement fundamentals. After we won approval for recruiting a measurement faculty member, the search and selection process taught us another important lesson - that the number of psychometricians with promising psychometric research programs interested in a faculty appointment rather than one with a psychometric research institute or testing company was small indeed.

Increasing the demand for measurement literacy and promoting the supply of measurement specialists

Our field is faced with a genuine supply and demand problem. The number of students entering measurement programs is limited by the number of available measurement faculty, and in turn the number of measurement courses and perception of need for measurement competence affects the ability to argue for faculty measurement positions. If one examines the faculty recruitment efforts in any academic year, the advertised measurement positions require a person who has hybrid credentials - measurement and statistics, or measurement and personality psychology, or measurement and social psychology. Nationwide, only a few of the major research universities have programs in measurement, but nearly all universities have programs for professional educators and psychologists.

There are several strategies for action for preserving measurement courses programs, and student bodies. One strategy is to capitalize on the

current popular methodologies such as "qualitative" methods and show how measurement theory can be applied to the technical problems encountered in rendering qualitative data into valid and reliable measures. The reform movement in education raises some other strategic possibilities. As professionals have shied away from tests with potential bias, and in turn moved toward an egalitarian or anti-individual difference value system, we now are faced with the challenge of developing new assessment methodologies to reflect cognitive and social research findings supporting interactive assessment, cognitive modifiability and social and/or cooperative learning. But to succeed in these efforts the measurement specialists will have maintain an involvement with current and future educational and psychological faculties as well as with future test users.

Proposals for increasing demand for measurement instruction and measurement literacy

The Joint Committee on Testing Practices is an interorganizational group representing APA, AERA, MCME and other organizations. It participates in various assessment and measurement projects, such as the Test User Qualifications Working Group. The recently published Code of Fair Testing Practices in Education is another example. The Joint Committee originated from a conference sponsored by APA in the Summer of 1984 in which test publishers and psychologists came together to develop a consensus on what the problems in testing were, and to develop an agreed upon approach to address particular problems in a cooperative way. One strategy for promoting measurement literacy is to support actively the implementation of the products that the Joint Committee develops both by working with publishers, as well as to provide reference sources for reviewing content of measurement courses in our university programs.

In addition to becoming aware of the potential of the several products of the Joint Committee on Testing Practices and utilizing them in our own instructional and consultation efforts, we can achieve a higher level of measurement literacy by reintegrating measurement faculty and measurement knowledge into the several reform movements in psychology and education.

Clarify methods for assessing program compliance with measurement standards in accreditation criteria. The APA accreditation standards are currently under revision, both to widen the scope of accreditation for clinical, counseling and school programs, and to provide for minimal as well as exemplary levels of program criteria. Regardless of the criteria for levels of accreditation that might be developed, it is reasonable to assume that instruction in measurement will continue to be required in every accredited psychology program approved by the APA accreditation process, including those that would be approved only at a minimal level. After all, testing and measurement has almost a 100 year history in APA and in psychology. If measurement psychology colleagues participate actively in developing standards for professional psychology programs, that include adequate measurement instruction, they will be furthering this longstanding history of measurement in psychology.

Division 5, 15, and other interested groups of psychologists can participate directly in this process by offering the technical assistance of measurement specialists who can assist in definitions and criteria for what would be considered to be a minimal level of instruction in measurement, and provide examples of ways that an accreditation review panel can evaluate a program's measurement offerings.

Participation in national conferences on graduate education and post-doctoral professional preparation for practice. In January, 1990, the Assembly of Scientist-Practicioner Psychologists will sponsor a national conference on the training of scientist-practitioner psychologists at the University of Florida in Gainesville. Division 5 and 15 can be sure that these concerns are represented in this conference. At the very least the conference should recommend ways that scientist-practitioner psychology programs can capitalize on new developments in measurement theory and incorporate them into their programs both in terms of faculty appointments, as well as in course instruction.

Reintegrate programs of measurement instruction into proposals for reform in the preparation of education professionals. In the proposals for reform of the ways that teachers and other education professionals are prepared for their practice roles, it is essential that measurement specialists, as

well as educational psychologists, contribute a perspective to these discussions that currently seems not to be present. The newsletters of the Holmes Group, one of the leading national organizations organized to consider reform in teacher education, composed of deans of schools of education, seem to reflect a perspective that centers attention on the conditions under which teachers teach, and ways to promote interest in teaching, and less on the knowledge base necessary for instruction. Some members of the Holmes Group are leaders in Division 5, and in educational and psychological measurement generally. Their attention needs to focus on the knowledge necessary for educational practice, and in the case of measurement literacy, the knowledge necessary to meet minimal standards for the appropriate selection, use, and interpretation of tests. We have a role to play here as well.

Review the content of national examinations for psychological and educational practitioners. The national examination used by all state psychology licensure boards as well as the national examinations for educational professionals can be reviewed to ascertain the extent to which measurement principles necessary for competent test use are integrated into the examinations. Taking into account the item writing competencies of many of our colleagues, we can expect that they can provide examples of measurement content can be proposed for inclusion in revisions of these examinations.

As new proposals for evaluating teacher competence by performance on teacher examinations are developed, raising the level of measurement literacy can be approached by incorporating items reflecting acceptable levels of test user competence.

Supply side possibilities

Create pre and postdoctoral fellowhips for promising measurement students. Having available pre and postdoctoral fellowships to be awarded to candidates in those institutions offering a graduate program in psychometrics might be a way of encouraging the interested students to select measurement as their chosen area of doctoral studies. Although some research and testing companies and State Boards of Education already have attempted this approach, combining the efforts of possible funding sources with

representative faculty at those instutitions with measurement programs, might create possibilities for improving the utilization of fellowships as a recuitment device, or to identify ways and places where the funds for student support might be best allocated. And in the event that those considering support of doctoral fellowships are concerned that the recipients might go on to careers in business rather than back to institutions or the measurement community, a payback provision could be introduced that would require the recipient to commit a term of service after graduation as a condition for receiving the fellowship award.

Endow chairs and increase the measurement supply. One strategy might be to negotiate with test publishers to set aside voluntarily a portion of their profits to endow chairs in psychometrics on university campuses. Since the qualifications of the person holding the chair can be determined in part by the endowment, test publishers and psychometric research institutes can have a direct impact on the future direction of measurement instruction and research.

Give prizes for outstanding examples of instructional modules to promote measurement literacy among teachers or practitioners. In my attempts to identify some of the problems associated with the demand and supply side of the measurement literacy matter, I do not intend to minimize the excellent work of professors and authors who have furthered this cause by their creation of excellent instructional materials. But formal recognition of these important applied measurement efforts might elevate their work to a platform where tangible and public recognition are available.

Encourage state boards of education to establish guidelines for testing specialists in schools. All efforts to improve the measurement literacy of teachers and educational practitioners should be supported by recognition of the need for school district employees and others who are responsible for the use and interpretation of tests in schools to have demonstrable competence in measurement. Negotiations with state school officers and school board presidents are the places to begin a direct attack of controlling the supply of literate measurement specialists in the public sector.

These proposals point out the many opportunities currently available for having a positive influence on the measurement competence of psychologists and educators. We can't expect test publishers to have sole responsibility for regulating test use by refusing to sell tests to users whose credentials do not reflect user competence. We can't expect employers to have sole responsibility for regulating test user competence by selecting only teachers, education professionals, and psychologists who reflect acceptable levels of user competence in measurement. And we can't expect that professionals will "make" themselves competent unless we in the measurement field assist them by defining minimal levels of measurement literacy necessary for various practice areas, by creating consultation resources that identify those who are willing to collaborate in reforming measurement knowledge levels and practices, and by identifying ways that measurement competence can be enhanced.

If you agree with me that the level of measurement literacy among educational and psychological professionals needs our immediate attention, then I hope that you will also agree that improving this state of affairs cannot be left to others. It is essentially our job since we are the ones with the knowledge that is not getting utilized.

Why Worry?

I contend that the future of measurement in psychology in the decades to come will be dependent on the availability of measurement faculty in our universities, the range of measurement offerings on our campuses, the standards for measurement literacy reflected in the preparation of psychological and educational professionals, in criteria for professional program accreditation, and standards for licenses and credentials for psychological and educational practice.

Our active participation in some of the activities that I have proposed above will affect directly the current level of supply and demand for psychometricians and psychometrics. We should acknowledge the excellent work of our colleagues who contributed to the revised Standards, and the several other testing documents that have been published subsequently as well as the current efforts of the Joint Committee on Testing Practices, and its subcommittees. But while they can develop the documents that support raising the level of measurement literacy among education and psychology professionals, the documents will not lead us far toward the goal of measurement literacy, unless we, as measurement folks, carry the movement back to our places of work, as well as outward to the various committees involved in proposals for revising national standards in these fields dependent on measurement literacy for competent practices.

References

Aiken, L. A., West, S. G., Sechrest, L. & Reno, R. R. (1989). *Graduate training in statistics, methodology, and measurement in psychology: A survey of Ph.D. program in North America.* Manuscript submitted for Publication.

American Association of State Psychology Board and the Professional Examination Service. (1984). *Examination for Professional Practice in Psychology* . New York: Professional Eamination Service.

American Psychological Association. (1979). *Criteria for accreditation of doctoral training programs and internships in professional psychology*. Washington, D. C. : Author.

American Psychological Association. (1985). *Standards for educational and psychological tests*. Washington, D. C. : Author.

Davison, M. L. Damarin, F. & Drasgow, F. (1986). Psychometrics and graduate education. *American Psychologist 42* 584-586.

Eyde, L., Moreland, Robertson, G. Primoff, & Most, R. (1988). Executive Summary. List User Qualification: A data-based approach to promoting good test use. Washington, D.C.: American Psychological Association.

Fremer, J., Diamond, E. & Cawara, W.J. (1989). Developing a code of fair testing practice in education. American Psychologist, 44 1062-1067.

Lambert, Robert E. (1980-81). Teacher attitudes on testing. A Multiple Perspective. *College Board Review* No. 118. 13,14,29,30.